

## Structural\_Data\_Types

The GEO-IDE effort acknowledges that NOAA's data systems are insufficiently integrated. This situation is a reflection of technology and management and decision-making strategies of the past that have tended to fragment data management, rather than to unify it. Lines of funding have traditionally been matched to observing system elements -- satellites, ships, profilers, etc.-- and data life cycle points -- measurement, real-time applications, climate analysis, archive, etc. Data management has been considered to be "owned" by the observing system element or the function. Each observing system element has therefore developed individualized approaches to data management, often involving the development of unique (and non-interoperable) data formats and protocols. Real-time data management strategies were devised with little thought to analysis or archive, and so on. Predictably these traditions have hindered the development integrated data management.

Communities of interest within data management are most naturally organized by structural type of data. The lines between these communities are drawn from the answers to key data management questions such as, What techniques are appropriate for searching for these data? For transporting (interchanging) these data? For visualizing or analyzing these data? For storing or archiving these data?

Communities of interest defined by structural data types provide a natural way to organize data management efforts and specify standards required for interoperability. For example, the kinds of standards, best practices, metadata, and access interfaces required for time-series data collections are similar for atmospheric, oceanic, hydrological, biological, or climate data.

Traditional communities of interest defined by pattern of usage will continue to thrive of course, based upon scientific and societal goals. These communities will provide the requirements to an increasingly integrated data management community. For example, weather forecasters will continue to require synoptic access to observations; climate modelers will continue to view the same observations as time series. The role of the data management community will be to find unified solutions that address both of these usage patterns.

Table 4.1 proposes an initial list of communities of interest based upon structural data types. In most cases the structural data types are the natural consequences of the manner in which the data are collected. For any given data stream there may be ambiguities regarding the appropriate structural data type under which it should be handled. As a general rule, the best way to resolve this ambiguity is to choose the most highly ordered data type that could describe the data. Table 4.1 is presented roughly in order from most highly structured data types at the top to least structured types at the bottom.

Structural Data Class	Descriptions and subclasses	Examples and further explanation
<b>Grids</b> (and collections of grids)	<ul style="list-style-type: none"><li>• <u>structured grids</u>: (uniform grids, rectilinear grids, curvilinear grids)</li><li>• <u>unstructured grids</u>: (grid cells defined by 3 or more nodes, linked by topology)</li></ul>	<ul style="list-style-type: none"><li>• finite difference model outputs (structured)</li><li>• finite element model outputs (unstructured)</li><li>• gridded (binned) data products (structured)</li><li>• level 4 (gridded) satellite fields (structured)</li><li>• spherical harmonic spectral coefficients (structured or unstructured)(1)</li></ul>
<b>Moving-sensor multidimensional fields</b> (and collections of same)	<ul style="list-style-type: none"><li>• swaths</li><li>• radials</li></ul>	<ul style="list-style-type: none"><li>• satellite passes</li><li>• HF radar</li><li>• side-scan sonar</li><li>• weather radar</li></ul>

## Structural\_Data\_Types

<b>Time series</b> (and collections of time series (2))	<ul style="list-style-type: none"> <li>time-ordered sequence of records (2) associated with a point in space or a more complex spatial feature.</li> </ul>	<ul style="list-style-type: none"> <li>ocean moored measurements (3)</li> <li>fish landings at a port</li> <li>stream flow records</li> <li>sun spot activity</li> <li>climate data (surface atmospheric stations)</li> <li>paleo-records from cores, corals, tree rings, ?</li> <li>computed climate indices such as SOI</li> </ul>
<b>Profiles</b> (and collections of profiles)	<ul style="list-style-type: none"> <li>height or depth-ordered sequence of records (1) at a fixed (or approximately fixed) point in time and position in lat/long</li> </ul>	<ul style="list-style-type: none"> <li>atmospheric soundings</li> <li>ocean casts</li> <li>profiling floats</li> <li>acoustic Doppler instruments (structural overlap with time series)</li> </ul>
<b>Trajectories</b> (and collections of trajectories)	<ul style="list-style-type: none"> <li>time-ordered sequence of records (2) along a path through space</li> </ul>	<ul style="list-style-type: none"> <li>underway ship measurements</li> <li>aircraft track data</li> <li>ocean surface drifters</li> <li>ocean AUV measurements</li> </ul>
<b>Geospatial Framework Data (4)</b>	<ul style="list-style-type: none"> <li>lines</li> <li>polygonal regions</li> <li>map annotations</li> </ul>	<ul style="list-style-type: none"> <li>shorelines</li> <li>fault lines</li> <li>marine boundaries</li> <li>continually operating reference stations (CORS)</li> </ul>
<b>Point data (5)</b>	<ul style="list-style-type: none"> <li>scattered points</li> </ul>	<ul style="list-style-type: none"> <li>tsunami or seismic occurrences</li> <li>species sitings</li> <li>geodetic control</li> <li>geospatial data</li> </ul>
<b>Metadata</b>	?data about data? ? context information needed for the interpretation of data	<p>Like other data types metadata has distinct requirements for storage, access, archival and transport.</p> <p>Metadata content is a major focus of discussions within all of the data types. Metadata as a ?data type? refers specifically to its unique requirement and properties with respect to archival, access, and transport.</p>

1. In some cases grids represent coordinate systems that are mathematically transformed from simple latitude-longitude-depth-time positions. Spherical harmonic spectral coefficients are an example of such.
2. A ?record? refers to one or more associated parameter values and associated metadata.
3. Standards for time series need to consider small, time-dependent excursions in latitude, longitude and depth. Cabled ocean moorings are an example of such.
4. The ?GIS perspective? must be a major focus in the discussion of all of the data classes listed in this table.
5. As an organizing principle for data ?Point Data? is the lowest common denominator. Most structural data types are reducible to collections of points, though with a loss of essential semantics in most

## Structural\_Data\_Types

cases. For example, a grid may be represented as a collection of ordered tuples. Some types of measurements, for example tsunami occurrences or species sightings, naturally possess limited structure. For these measurements the Point Data structure is the natural classification.

Note that real time delivery of data will generally remove time structure, so that, for example, a collection of Time Series may reduce to Point Data when accessed in real time.

Global Earth Observation Integrated Data Environment CONOPS